SEQUENCE LISTING

<110> Chen, Fang

<120> DNA MOLECULES ENCODING HUMAN NUCLEAR RECEPTOR PROTEINS

<130> 19999Y

<160> 30

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 2807
<212> DNA
<213> Human

<400> 1

```
gaatatgatg accctaatgc aacaatatct aacatactat ccgagcttcg gtcatttgga      60
agaactgcag attttcctcc ttcaaaatta aagtcaggtt atggagaaca tgtatgctat     120
gttcttgatt gcttcgctga agaagcattg aaatatattg gtttcacctg gaaaaggcca     180
atatacccag tagaagaatt agaagaagaa agcgttgcag aagatgatgc agaattaaca     240
ttaaataaag tggatgaaga atttgtggaa gaagagacag ataatgaaga aaactttatt     300
gatctcaacg ttttaaaggc ccagacatat cacttggata tgaacgagac tgccaaacaa     360
gaagatattt tggaatccac aacagatgct gcagaatgga gcctagaagt ggaacgtgta     420
ctaccgcaac tgaaagtcac gattaggact gacaataagg attggagaat ccatgttgac     480
caaatgcacc agcacagaag tggaattgaa tctgctctaa aggagaccaa gggatttttg     540
gacaaactcc ataatgaaat tactaggact ttggaaaaga tcagcagccg agaaaagtac     600
atcaacaatc agccgggagc ccatggagca ctgtcctcag agatgcgcag gttaggctca     660
ctgtctaggc caggcccacc ttagtcactg tggactggca atggaagctc ttcctggaca     720
cacctgccct agccctcacc ctggggtgga agagaaatga gcttggcttg caactcagac     780
cattccacgg aggcatcctc cccttccctg ggctggtgaa taaaagtttc ctgaggtcaa     840
ggacttcctt ttccctgcca aaatggtgtc cagaactttg aggccagagg tgatccagtg     900
atttgggagc tgcaggtcac acaggctgct cagagggctg ctgaacagga tgtcctcgga     960
cgacaggcac ctgggctcca gctgcggctc cttcatcaag actgagccgt ccagcccgtc    1020
ctcgggcata gatgccctca gccaccacag ccccagtggc tcgtccgacg ccagcggcgg    1080
ctttggcctg gccctgggca cccacgccaa cggtctggac tcgccaccca tgtttgcagg    1140
cgccgggctg ggaggcaccc catgccgcaa gagctacgag gactgtgcca gcggcatcat    1200
ggaggactcg gccatcaagt gcgagtacat gctcaacgcc atccccaagc gcctgtgcct    1260
cgtgtgcggg gacattgcct ctggctacca ctacggcgtg gcctcctgcg aggcttgcaa    1320
ggccttcttc aagaggacta tccaagggaa cattgagtac agctgcccgg ccaccaacga    1380
gtgcgagatc accaaacgga ggcgcaagtc ctgccaggcc tgccgcttca tgaaatgcct    1440
caaagtgggg atgctgaagg aaggtgtgcg ccttgatcga gtgcgtggag gccgtcagaa    1500
atacaagcga cggctggact cagagagcag cccatacctg agcttacaaa tttctccacc    1560
tgctaaaaag ccattgacca agattgtctc atacctactg gtggctgagc cggacaagct    1620
ctatgccatg cctccccctg gtatgcctga gggggacatc aaggccctga ccactctctg    1680
tgacctggca gaccgagagc ttgtggtcat cattggctgg gccaagcaca tcccaggctt    1740
ctcaagcctc tccctggggg accagatgag cctgctgcag agtgcctgga tggaaatcct    1800
catcctgggc atcgtgtacc gctcgctgcc ctacgacgac aagctggtgt acgctgagga    1860
ctacatcatg gatgaggagc actcccgcct cgcggggctg ctggagctct accgggccat    1920
cctgcagctg gtacgcaggt acaagaagct caaggtggag aaggaggagt ttgtgacgct    1980
caaggccctg gccctcgcca actccgattc catgtacatc gaggatctag aggctgtcca    2040
gaagctgcag gacctgctgc acgaggcact gcaggactac gagctgagcc agcgccatga    2100
ggagccctgg aggacgggca agctgctgct gacactgccg ctgctgcggc agacggccgc    2160
```

caaggccgtg cagcacttct atagcgtcaa actgcagggc aaagtgccca tgcacaaact 2220
cttcctggag atgctggagg ccaaggcctg ggccagggct gactcccttc aggagtggag 2280
gccactggag caagtgccct ctcccctcca ccgagccacc aagaggcagc atgtgcattt 2340
cctaactccc ttgccccctc ccccatctgt ggcctgggtg ggcactgctc aggctggata 2400
ccacctggag gttttccttc cgcagagggc aggttggcca agagcagctt agaggatctc 2460
ccaaggatga aagaatgtca agccatgatg gaaaatgccc cttccaatca gctgccttca 2520
caagcaggga tcagagcaac tccccgggga tccccaatcc acgcccttct agtccaaccc 2580
ccctcaatga gagaggcagg cagatctcac ccagcactag gacaccagga ggccagggaa 2640
agcatctctg gctcaccatg taacatctgg cttggagcaa gtgggtgttc tgcacaccag 2700
gcagctgcac ctcactggat ctagtgttgc tgcgagtgac ctcacttcag agcccctcta 2760
gcagagtggg gcggaagtcc tgatggttgg tgtccatgag gtggaag 2807

<210> 2
<211> 500
<212> PRT
<213> Human

<400> 2

Met Ser Ser Asp Asp Arg His Leu Gly Ser Ser Cys Gly Ser Phe Ile
1 5 10 15
Lys Thr Glu Pro Ser Ser Pro Ser Ser Gly Ile Asp Ala Leu Ser His
20 25 30
His Ser Pro Ser Gly Ser Ser Asp Ala Ser Gly Gly Phe Gly Leu Ala
35 40 45
Leu Gly Thr His Ala Asn Gly Leu Asp Ser Pro Pro Met Phe Ala Gly
50 55 60
Ala Gly Leu Gly Gly Thr Pro Cys Arg Lys Ser Tyr Glu Asp Cys Ala
65 70 75 80
Ser Gly Ile Met Glu Asp Ser Ala Ile Lys Cys Glu Tyr Met Leu Asn
85 90 95
Ala Ile Pro Lys Arg Leu Cys Leu Val Cys Gly Asp Ile Ala Ser Gly
100 105 110
Tyr His Tyr Gly Val Ala Ser Cys Glu Ala Cys Lys Ala Phe Phe Lys
115 120 125
Arg Thr Ile Gln Gly Asn Ile Glu Tyr Ser Cys Pro Ala Thr Asn Glu
130 135 140
Cys Glu Ile Thr Lys Arg Arg Arg Lys Ser Cys Gln Ala Cys Arg Phe
145 150 155 160
Met Lys Cys Leu Lys Val Gly Met Leu Lys Glu Gly Val Arg Leu Asp
165 170 175
Arg Val Arg Gly Gly Arg Gln Lys Tyr Lys Arg Arg Leu Asp Ser Glu
180 185 190
Ser Ser Pro Tyr Leu Ser Leu Gln Ile Ser Pro Pro Ala Lys Lys Pro
195 200 205
Leu Thr Lys Ile Val Ser Tyr Leu Leu Val Ala Glu Pro Asp Lys Leu
210 215 220
Tyr Ala Met Pro Pro Pro Gly Met Pro Glu Gly Asp Ile Lys Ala Leu
225 230 235 240
Thr Thr Leu Cys Asp Leu Ala Asp Arg Glu Leu Val Val Ile Ile Gly
245 250 255
Trp Ala Lys His Ile Pro Gly Phe Ser Ser Leu Ser Leu Gly Asp Gln
260 265 270
Met Ser Leu Leu Gln Ser Ala Trp Met Glu Ile Leu Ile Leu Gly Ile
275 280 285
Val Tyr Arg Ser Leu Pro Tyr Asp Asp Lys Leu Val Tyr Ala Glu Asp
290 295 300

```
Tyr Ile Met Asp Glu Glu His Ser Arg Leu Ala Gly Leu Leu Glu Leu
305                 310                 315                 320
Tyr Arg Ala Ile Leu Gln Leu Val Arg Arg Tyr Lys Lys Leu Lys Val
                325                 330                 335
Glu Lys Glu Glu Phe Val Thr Leu Lys Ala Leu Ala Leu Ala Asn Ser
            340                 345                 350
Asp Ser Met Tyr Ile Glu Asp Leu Glu Ala Val Gln Lys Leu Gln Asp
        355                 360                 365
Leu Leu His Glu Ala Leu Gln Asp Tyr Glu Leu Ser Gln Arg His Glu
    370                 375                 380
Glu Pro Trp Arg Thr Gly Lys Leu Leu Leu Thr Leu Pro Leu Leu Arg
385                 390                 395                 400
Gln Thr Ala Ala Lys Ala Val Gln His Phe Tyr Ser Val Lys Leu Gln
                405                 410                 415
Gly Lys Val Pro Met His Lys Leu Phe Leu Glu Met Leu Glu Ala Lys
            420                 425                 430
Ala Trp Ala Arg Ala Asp Ser Leu Gln Glu Trp Arg Pro Leu Glu Gln
        435                 440                 445
Val Pro Ser Pro Leu His Arg Ala Thr Lys Arg Gln His Val His Phe
    450                 455                 460
Leu Thr Pro Leu Pro Pro Pro Pro Ser Val Ala Trp Val Gly Thr Ala
465                 470                 475                 480
Gln Ala Gly Tyr His Leu Glu Val Phe Leu Pro Gln Arg Ala Gly Trp
                485                 490                 495
Pro Arg Ala Ala
            500
```

<210> 3
<211> 2985
<212> DNA
<213> Human

<400> 3

```
gcgggccgcc agtgtggtgg aattcggctt gtcactagga gaacatttgt gttaattgca      60
ctgtgctctg tcaaggaaac tttgatttat agctggggtg cacaaataat ggttgccggt     120
cgcacatgga ttcggtagaa ctttgccttc ctgaatcttt ttccctgcac tacgaggaag     180
agcttctctg cagaatgtca aacaaagatc gacacattga ttccagctgt tcgtccttca     240
tcaagacgga accttccagc ccagcctccc tgacggacag cgtcaaccac cacagccctg     300
gtggctcttc agacgccagt gggagctaca gttcaaccat gaatggccat cagaacggac     360
ttgactcgcc acctctctac ccttctgctc ctatcctggg aggtagtggg cctgtcagga     420
aactgtatga tgactgctcc agcaccattg ttgaagatcc ccagaccaag tgtgaataca     480
tgctcaactc gatgcccaag agactgtgtt tagtgtgtgg tgacatcgct tctgggtacc     540
actatggggt agcatcatgt gaagcctgca aggcattctt caagaggaca attcaaggca     600
atatagaata cagctgccct gccacgaatg aatgtgaaat cacaaagcgc agacgtaaat     660
cctgccaggc ttgccgcttc atgaagtgtt taaaagtggg catgctgaaa gaaggggtgc     720
gtcttgacag agtacgtgga ggtcggcaga agtacaagcg caggatagat gcggagaaca     780
gcccatacct gaaccctcag ctggttcagc cagccaaaaa gccatataac aagattgtct     840
cacatttgtt ggtggctgaa ccggagaaga tctatgccat gcctgaccct actgtccccg     900
acagtgacat caaagccctc actacactgt gtgacttggc cgaccgagag ttggtggtta     960
tcattggatg ggcgaagcat attccaggct tctccacgct gtccctggcg gaccagatga    1020
gccttctgca gagtgcttgg atggaaattt tgatccttgg tgtcgtatac cggtctcttt    1080
catttgagga tgaacttgtc tatgcagacg attatataat ggacgaagac cagtccaaat    1140
tagcaggcct tcttgatcta aataatgcta tcctgcagct ggtaaagaaa tacaagagca    1200
tgaagctgga aaaagaagaa tttgtcaccc tcaaagctat agctcttgct aattcagact    1260
ccatgcacat agaagatgtt gaagccgttc agaagcttca ggatgtctta catgaagcgc    1320
tgcaggatta tgaagctggc cagcacatgg aagaccctcg tcgagctggc aagatgctga    1380
```

```
tgacactgcc actcctgagg cagacctcta ccaaggccgt gcagcatttc tacaacatca    1440
aactagaagg caaagtccca atgcacaaac ttttttgga aatgttggag gccaaggtct    1500
gactaaaagc tccctgggcc ttcccatcct tcatgttgaa aaagggaaaa taaacccaag    1560
agtgatgtcg aagaaactta gagtttagtt aacaacatca aaaatcaaca gactgcactg    1620
ataatttagc agcaagacta tgaagcagct ttcagattcc tccataggtt cctgatgagt    1680
tctttctact ttctccatca tcttctttcc tctttcttcc cacatttctc tttctcttta    1740
ttttttctcc ttttcttctt tcacctccct tatttctttg cttctttcat tcctagttcc    1800
cattctcctt tatttttcttc ccgtctgcct gccttctttc ttttctttac ctactctcat    1860
tcctctcttt tctcatcctt ccccttttt ctaaatttga aatagcttta gtttaaaaaa    1920
aaaaatcctc ccttcccccc ttcctttccc tttctttcct ttttcccttt ccttttccct    1980
ttcctttcct ttcctcttga ccttctttcc atctttcttt ttcttccttc tgctgctgaa    2040
cttttaaaag aggtctctaa ctgaagagag atggaagcca gccctgccaa aggatggaga    2100
tccataatat ggatgccagt gaacttattg tgaaccatac cgtccccaat gactaaggaa    2160
tcaaagagag agaaccaacg ttcctaaaag tacagtgcaa catatacaaa ttgactgagt    2220
gcagtattag atttcatggg agcagcctct aattagacaa cttaagcaac gttgcatcgg    2280
ctgcttctta tcattgcttt tccatctaga tcagttacag ccatttgatt ccttaattgt    2340
tttttcaagt cttccaggta tttgttagtt tagctactat gtaacttttt cagggaatag    2400
tttaagcttt attcattcat gcaatactaa agagaaataa gaatactgca attttgtgct    2460
ggctttgaac aattacgaac aataatgaag gacaaatgaa tcctgaagga agatttttaa    2520
aaatgttttg tttcttctta caaatggaga tttttttgta ccagctttac cacttttcag    2580
ccatttatta atatgggaat ttaacttact caagcaatag ttgaagggaa ggtgcatatt    2640
atcacggatg caatttatgt tgtgtgccag tctggtccca aacatcaatt tcttaacatg    2700
agctccagtt tacctaaatg ttcactgaca caaaggatga gattacacct acagtgactc    2760
tgagtagtca catatataag cactgcacat gagatataga tccgtagaat tgtcaggagt    2820
gcacctctct acttgggagg tacaattgcc atatgatttc tagctgccat ggtggttagg    2880
aatgtgatac tgcctgtttg caaagtcaca gaccttgcct cagaaggagc tgtgagccag    2940
tattcattta agagaattcc accacactgg cggcccgcgc ttgat                    2985
```

<210> 4
<211> 458
<212> PRT
<213> Human

<400> 4

```
Met Asp Ser Val Glu Leu Cys Leu Pro Glu Ser Phe Ser Leu His Tyr
 1               5                  10                  15
Glu Glu Glu Leu Leu Cys Arg Met Ser Asn Lys Asp Arg His Ile Asp
            20                  25                  30
Ser Ser Cys Ser Ser Phe Ile Lys Thr Glu Pro Ser Ser Pro Ala Ser
        35                  40                  45
Leu Thr Asp Ser Val Asn His His Ser Pro Gly Gly Ser Ser Asp Ala
    50                  55                  60
Ser Gly Ser Tyr Ser Ser Thr Met Asn Gly His Gln Asn Gly Leu Asp
65                  70                  75                  80
Ser Pro Pro Leu Tyr Pro Ser Ala Pro Ile Leu Gly Gly Ser Gly Pro
                85                  90                  95
Val Arg Lys Leu Tyr Asp Asp Cys Ser Ser Thr Ile Val Glu Asp Pro
            100                 105                 110
Gln Thr Lys Cys Glu Tyr Met Leu Asn Ser Met Pro Lys Arg Leu Cys
        115                 120                 125
Leu Val Cys Gly Asp Ile Ala Ser Gly Tyr His Tyr Gly Val Ala Ser
    130                 135                 140
Cys Glu Ala Cys Lys Ala Phe Phe Lys Arg Thr Ile Gln Gly Asn Ile
145                 150                 155                 160
Glu Tyr Ser Cys Pro Ala Thr Asn Glu Cys Glu Ile Thr Lys Arg Arg
                165                 170                 175
```

```
Arg Lys Ser Cys Gln Ala Cys Arg Phe Met Lys Cys Leu Lys Val Gly
            180                 185                 190
Met Leu Lys Glu Gly Val Arg Leu Asp Arg Val Arg Gly Gly Arg Gln
        195                 200                 205
Lys Tyr Lys Arg Arg Ile Asp Ala Glu Asn Ser Pro Tyr Leu Asn Pro
    210                 215                 220
Gln Leu Val Gln Pro Ala Lys Lys Pro Tyr Asn Lys Ile Val Ser His
225                 230                 235                 240
Leu Leu Val Ala Glu Pro Glu Lys Ile Tyr Ala Met Pro Asp Pro Thr
                245                 250                 255
Val Pro Asp Ser Asp Ile Lys Ala Leu Thr Thr Leu Cys Asp Leu Ala
            260                 265                 270
Asp Arg Glu Leu Val Val Ile Ile Gly Trp Ala Lys His Ile Pro Gly
        275                 280                 285
Phe Ser Thr Leu Ser Leu Ala Asp Gln Met Ser Leu Leu Gln Ser Ala
    290                 295                 300
Trp Met Glu Ile Leu Ile Leu Gly Val Val Tyr Arg Ser Leu Ser Phe
305                 310                 315                 320
Glu Asp Glu Leu Val Tyr Ala Asp Asp Tyr Ile Met Asp Glu Asp Gln
                325                 330                 335
Ser Lys Leu Ala Gly Leu Leu Asp Leu Asn Asn Ala Ile Leu Gln Leu
            340                 345                 350
Val Lys Lys Tyr Lys Ser Met Lys Leu Glu Lys Glu Glu Phe Val Thr
        355                 360                 365
Leu Lys Ala Ile Ala Leu Ala Asn Ser Asp Ser Met His Ile Glu Asp
    370                 375                 380
Val Glu Ala Val Gln Lys Leu Gln Asp Val Leu His Glu Ala Leu Gln
385                 390                 395                 400
Asp Tyr Glu Ala Gly Gln His Met Glu Asp Pro Arg Arg Ala Gly Lys
                405                 410                 415
Met Leu Met Thr Leu Pro Leu Leu Arg Gln Thr Ser Thr Lys Ala Val
            420                 425                 430
Gln His Phe Tyr Asn Ile Lys Leu Glu Gly Lys Val Pro Met His Lys
        435                 440                 445
Leu Phe Leu Glu Met Leu Glu Ala Lys Val
    450                 455
```

<210> 5
<211> 2987
<212> DNA
<213> Human

<400> 5

```
gcgggccgcc agtgtggtgg aattcggctt gtcactagga gaacatttgt gttaattgca      60
ctgtgctctg tcaaggaaac tttgatttat agctggggtg cacaaataat ggttgccggt     120
cgcacatgga ttcggtagaa ctttgccttc ctgaatcttt ttccctgcac tacgaggaag     180
agcttctctg cagaatgtca aacaaagatc gacacattga ttccagctgt tcgtccttca     240
tcaagacgga accttccagc ccagcctccc tgacggacag cgtcaaccac cacagccctg     300
gtggctcttc agacgccagt gggagctaca gttcaaccat gaatggccat cagaacggac     360
ttgactcgcc acctctctac ccttctgctc ctatcctggg aggtagtggg cctgtcagga     420
aactgtatga tgactgctcc agcaccattg ttgaagatcc ccagaccaag tgtgaataca     480
tgctcaactc gatgcccaag agactgtgtt tagtgtgtgg tgacatcgct tctgggtacc     540
actatggggt agcatcatgt gaagcctgca aggcattctt caagaggaca attcaaggca     600
atatagaata cagctgccct gccacgaatg aatgtgaaat cacaaagcgc agacgtaaat     660
cctgccaggc ttgccgcttc atgaagtgtt taaaagtggg catgctgaaa gaaggggtgc     720
gtcttgacag agtacgtgga ggtcggcaga agtacaagcg caggatagat gcggagaaca     780
```

```
gcccatacct gaaccctcag ctggttcagc cagccaaaaa gccatataac aagattgtct     840
cacatttgtt ggtggctgaa ccggagaaga tctatgccat gcctgaccct actgtccccg     900
acagtgacat caaagccctc actacactgt gtgacttggc cgaccgagag ttggtggtta     960
tcattggatg ggcgaagcat attccaggct tctccacgct gtccctggcg gaccagatga    1020
gccttctgca gagtgcttgg atggaaattt tgatccttgg tgtcgtatac cggtctcttt    1080
catttgagga tgaacttgtc tatgcagacg attatataat ggacgaagac cagtccaaat    1140
tagcaggcct tcttgatcta aataatgcta tcctgcagct ggtaaagaaa tacaagagca    1200
tgaagctgga aaaagaagaa tttgtcaccc tcaaagctat agctcttgct aattcagact    1260
ccatgcacat agaagatgtt gaagccgttc agaagcttca ggatgtctta catgaagcgc    1320
tgcaggatta tgaagctggc cagcacatgg agaagaccct cgtcgagctg gcaagatgct    1380
gatgacactg ccactcctga ggcagacctc taccaaggcc gtgcagcatt tctacaacat    1440
caaactagaa ggcaaagtcc caatgcacaa acttttttg gaaatgttgg aggccaaggt     1500
ctgactaaaa gctccctggg ccttcccatc cttcatgttg aaaaagggaa aataaaccca    1560
agagtgatgt cgaagaaact tagagtttag ttaacaacat caaaaatcaa cagactgcac    1620
tgataattta gcagcaagac tatgaagcag ctttcagatt cctccatagg ttcctgatga    1680
gttctttcta ctttctccat catcttcttt cctctttctt cccacatttc tctttctctt    1740
tatttttct cctttcttc tttcacctcc cttatttctt tgcttctttc attcctagtt     1800
cccattctcc tttatttttct tcccgtctgc ctgccttctt tcttttcttt acctactctc    1860
attcctctct tttctcatcc ttccccttt ttctaaattt gaaatagctt tagtttaaaa     1920
aaaaaaatcc tcccttcccc ctttcctttc cctttctttc ctttttccct ttcctttccc    1980
ctttcctttc ctttcctctt gaccttcttt ccatctttct ttttcttcct tctgctgctg    2040
aacttttaaa agaggtctct aactgaagag agatggaagc cagccctgcc aaaggatgga    2100
gatccataat atggatgcca gtgaacttat tgtgaaccat accgtcccca atgactaagg    2160
aatcaaagag agagaaccaa cgttcctaaa agtacagtgc aacatataca aattgactga    2220
gtgcagtatt agatttcatg ggagcagcct ctaattagac aacttaagca acgttgcatc    2280
ggctgcttct tatcattgct tttccatcta gatcagttac agccatttga ttccttaatt    2340
gtttttttcaa gtcttccagg tatttgttag tttagctact atgtaacttt ttcagggaat    2400
agtttaagct ttattcattc atgcaatact aaagagaaat aagaatactg caattttgtg    2460
ctggctttga acaattacga acaataatga aggacaaatg aatcctgaag gaagattttt    2520
aaaaatgttt tgtttcttct tacaaatgga gatttttttg taccagcttt accacttttc    2580
agccatttat taatatggga atttaactta ctcaagcaat agttgaaggg aaggtgcata    2640
ttatcacgga tgcaatttat gttgtgtgcc agtctggtcc caaacatcaa tttcttaaca    2700
tgagctccag tttacctaaa tgttcactga cacaaaggat gagattacac ctacagtgac    2760
tctgagtagt cacatatata agcactgcac atgagatata gatccgtaga attgtcagga    2820
gtgcacctct ctacttggga ggtacaattg ccatatgatt tctagctgcc atggtggtta    2880
ggaatgtgat actgcctgtt tgcaaagtca cagaccttgc ctcagaagga gctgtgagcc    2940
agtattcatt taagagaatt ccaccacact ggcggcccgc gcttgat                  2987
```

<210> 6
<211> 418
<212> PRT
<213> Human

<400> 6

```
Met Asp Ser Val Glu Leu Cys Leu Pro Glu Ser Phe Ser Leu His Tyr
 1               5                  10                  15
Glu Glu Glu Leu Leu Cys Arg Met Ser Asn Lys Asp Arg His Ile Asp
            20                  25                  30
Ser Ser Cys Ser Ser Phe Ile Lys Thr Glu Pro Ser Ser Pro Ala Ser
        35                  40                  45
Leu Thr Asp Ser Val Asn His His Ser Pro Gly Gly Ser Ser Asp Ala
    50                  55                  60
Ser Gly Ser Tyr Ser Ser Thr Met Asn Gly His Gln Asn Gly Leu Asp
65                  70                  75                  80
Ser Pro Pro Leu Tyr Pro Ser Ala Pro Ile Leu Gly Gly Ser Gly Pro
                85                  90                  95
```

Val Arg Lys Leu Tyr Asp Asp Cys Ser Ser Thr Ile Val Glu Asp Pro
100 105 110
Gln Thr Lys Cys Glu Tyr Met Leu Asn Ser Met Pro Lys Arg Leu Cys
115 120 125
Leu Val Cys Gly Asp Ile Ala Ser Gly Tyr His Tyr Gly Val Ala Ser
130 135 140
Cys Glu Ala Cys Lys Ala Phe Phe Lys Arg Thr Ile Gln Gly Asn Ile
145 150 155 160
Glu Tyr Ser Cys Pro Ala Thr Asn Glu Cys Glu Ile Thr Lys Arg Arg
165 170 175
Arg Lys Ser Cys Gln Ala Cys Arg Phe Met Lys Cys Leu Lys Val Gly
180 185 190
Met Leu Lys Glu Gly Val Arg Leu Asp Arg Val Arg Gly Gly Arg Gln
195 200 205
Lys Tyr Lys Arg Arg Ile Asp Ala Glu Asn Ser Pro Tyr Leu Asn Pro
210 215 220
Gln Leu Val Gln Pro Ala Lys Lys Pro Tyr Asn Lys Ile Val Ser His
225 230 235 240
Leu Leu Val Ala Glu Pro Glu Lys Ile Tyr Ala Met Pro Asp Pro Thr
245 250 255
Val Pro Asp Ser Asp Ile Lys Ala Leu Thr Thr Leu Cys Asp Leu Ala
260 265 270
Asp Arg Glu Leu Val Val Ile Ile Gly Trp Ala Lys His Ile Pro Gly
275 280 285
Phe Ser Thr Leu Ser Leu Ala Asp Gln Met Ser Leu Leu Gln Ser Ala
290 295 300
Trp Met Glu Ile Leu Ile Leu Gly Val Val Tyr Arg Ser Leu Ser Phe
305 310 315 320
Glu Asp Glu Leu Val Tyr Ala Asp Asp Tyr Ile Met Asp Glu Asp Gln
325 330 335
Ser Lys Leu Ala Gly Leu Leu Asp Leu Asn Asn Ala Ile Leu Gln Leu
340 345 350
Val Lys Lys Tyr Lys Ser Met Lys Leu Glu Lys Glu Glu Phe Val Thr
355 360 365
Leu Lys Ala Ile Ala Leu Ala Asn Ser Asp Ser Met His Ile Glu Asp
370 375 380
Val Glu Ala Val Gln Lys Leu Gln Asp Val Leu His Glu Ala Leu Gln
385 390 395 400
Asp Tyr Glu Ala Gly Gln His Met Glu Lys Thr Leu Val Glu Leu Ala
405 410 415
Arg Cys

<210> 7
<211> 403
<212> DNA
<213> Human

<400> 7
cttttttagga ggtggagaaa tttgtaagct caggtatggg ctgctctctg agtccagccg 60
tcgcttgtat ttctgacggc ctccacgcac tcgatcaagg cgcacacctt ccttcagcat 120
ccccactttg aggcatttca tgaagcggca ggcctggcag gacttgcgcc tccgtttggt 180
gatctcgcac tcgttggtgg ccgggcagct gtactcaatg ttcccttgga tagtcctctt 240
gaagaaggcc ttgcaagcct cgcaggaggc ccacgcgtna gtggtagcca gagnaaatgt 300
ccccgcacac gaggcacagg cgcttgggga tggcgttgag catgttactt cgcacttgga 360
tgggccgagt cctccatgga tggccgctgg caacagttcc tcg 403

<210> 8
<211> 622
<212> DNA
<213> Human

<400> 8
cnnnnnnnnn nnntttttnnt gcctaaagtg gtacccngaa gngatgtcac cacacactaa 60
acacagtctc ttgggcatcg agttgagcat gtattcacac ttggtctggg gatcttcaac 120
aatggtgctg gagcagtcat catacagttt cctgacaggc ccactacctc ccaggatagg 180
agcagaaggg tagagaggtg gcgagtcaag tccgttctga tggccattca tggttgaact 240
gtagctccca ctggcgtctg aagagccacc agggctgtgg tggttgacgc tgtccgtcag 300
ggaggctggg ctggaaggtt ccgtcttgat gaaggacgaa cagctggaat caatgtgtcg 360
atctttgttt ggacattctg cagagaagct cttcctccgt ngtgcaggga aaaagattca 420
ggaaggcaaa gttcttcccg aatccatgtg cgaccggaaa ccattatttg ngcaccccag 480
ctattaatca aagttccttg acagagacag ggcaattaca naatgtctcc tntnggggat 540
caactgttcn gtattnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 600
nnnnnnnnnn nnnnnnnnnn tt 622

<210> 9
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 9
tgagtccagc cgtcgcttgt at 22

<210> 10
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 10
tgcaagcctc gcaggaggcc 20

<210> 11
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 11
ggccttcttc aagaggacta tc 22

<210> 12
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 12
aaagatcgac acattgattc c 21

<210> 13
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 13
gacttgactc gccacctctc 20

<210> 14
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 14
gttctgatgg ccattcatgg t 21

<210> 15
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 15
gaatatgatg accctaatgc a 21

<210> 16
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 16
cttccacctc atggacacca a 21

<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 17
gttaattgca ctgtgctctg 20

<210> 18
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 18
agtgtggtgg aattctctta 20

<210> 19
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 19
tctagtgttg ctgcgagtga c 21

<210> 20
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 20
cttccacctc atggacacca a 21

<210> 21
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 21
gtctgactaa aagctccctg 20

<210> 22
<211> 21
<212> DNA
<213> Artificial Sequence

<220>

<223> oligonucleotide

<400> 22
gaagatgatg gagaaagtag a 21

<210> 23
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 23
cattccacgg aggcatcctc 20

<210> 24
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 24
ccaaggccgt gcagcacttc 20

<210> 25
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 25
gacagcctct agatcctcga t 21

<210> 26
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 26
atcatggctt gacattcttt c 21

<210> 27
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 27
agctcttgct aattcagac 19

<210> 28
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<400> 28
tcaacatgaa ggatgggaag g 21

<210> 29
<211> 2807
<212> DNA
<213> Human

<400> 29
cttatactac tgggattacg ttgttataga ttgtatgata ggctcgaagc cagtaaacct 60
tcttgacgtc taaaaggagg aagttttaat ttcagtccaa tacctcttgt acatacgata 120
caagaactaa cgaagcgact tcttcgtaac tttatataac caaagtggac cttttccggt 180
tatatgggtc atcttcttaa tcttcttctt tcgcaacgtc ttctactacg tcttaattgt 240
aatttatttc acctacttct taaacacctt cttctctgtc tattacttct tttgaaataa 300
ctagagttgc aaaatttccg ggtctgtata gtgaacctat acttgctctg acggtttgtt 360
cttctataaa accttaggtg ttgtctacga cgtcttacct cggatcttca ccttgcacat 420
gatggcgttg actttcagtg ctaatcctga ctgttattcc taacctctta ggtacaactg 480
gtttacgtgg tcgtgtcttc accttaactt agacgagatt tcctctggtt ccctaaaaac 540
ctgtttgagg tattacttta atgatcctga aacctttttct agtcgtcggc tcttttcatg 600
tagttgttag tcggccctcg ggtacctcgt gacaggagtc tctacgcgtc caatccgagt 660
gacagatccg gtccgggtgg aatcagtgac acctgaccgt taccttcgag aaggacctgt 720
gtggacggga tcgggagtgg gaccccacct tctctttact cgaaccgaac gttgagtctg 780
gtaaggtgcc tccgtaggag gggaagggac ccgaccactt attttcaaag gactccagtt 840
cctgaaggaa aagggacggt tttaccacag gtcttgaaac tccggtctcc actaggtcac 900
taaaccctcg acgtccagtg tgtccgacga gtctcccgac gacttgtcct acaggagcct 960
gctgtccgtg gacccgaggt cgacgccgag gaagtagttc tgactcggca ggtcgggcag 1020
gagcccgtat ctacgggagt cggtggtgtc ggggtcaccg agcaggctgc ggtcgccgcc 1080
gaaaccggac cgggacccgt gggtgcggtt gccagacctg agcggtgggt acaaacgtcc 1140
gcggcccgac cctccgtggg gtacggcgtt ctcgatgctc ctgacacggt cgccgtagta 1200
cctcctgagc cggtagttca cgctcatgta cgagttgcgg taggggttcg cggacacgga 1260
gcacacgccc ctgtaacgga gaccgatggt gatgccgcac cggaggacgc tccgaacgtt 1320
ccggaagaag ttctcctgat aggttccctt gtaactcatg tcgacgggcc ggtggttgct 1380
cacgctctag tggtttgcct ccgcgttcag gacggtccgg acggcgaagt actttacgga 1440
gtttcacccc tacgacttcc ttccacacgc ggaactagct cacgcacctc cggcagtctt 1500
tatgttcgct gccgacctga gtctctcgtc gggtatggac tcgaatgttt aaagaggtgg 1560
acgatttttc ggtaactggt tctaacagag tatggatgac caccgactcg gcctgttcga 1620
gatacggtac ggaggggggac catacggact ccccctgtag ttccgggact ggtgagagac 1680
actggaccgt ctggctctcg aacaccagta gtaaccgacc cggttcgtgt agggtccgaa 1740
gagttcggag agggaccccc tggtctactc ggacgacgtc tcacggacct acctttagga 1800
gtaggacccg tagcacatgg cgagcgacgg gatgctgctg ttcgaccaca tgcgactcct 1860
gatgtagtac ctactcctcg tgagggcgga gcgccccgac gacctcgaga tggcccggta 1920
ggacgtcgac catgcgtcca tgttcttcga gttccacctc ttcctcctca aacactgcga 1980
gttccgggac cgggagcggt tgaggctaag gtacatgtag ctcctagatc tccgacaggt 2040
cttcgacgtc ctggacgacg tgctccgtga cgtcctgatg ctcgactcgg tcgcggtact 2100

cctcgggacc tcctgcccgt tcgacgacga ctgtgacggc gacgacgccg tctgccggcg 2160
gttccggcac gtcgtgaaga tatcgcagtt tgacgtcccg tttcacgggt acgtgtttga 2220
gaaggacctc tacgacctcc ggttccggac ccggtcccga ctgagggaag tcctcacctc 2280
cggtgacctc gttcacggga gaggggaggt ggctcggtgg ttctccgtcg tacacgtaaa 2340
ggattgaggg aacgggggag ggggtagaca ccggacccac ccgtgacgag tccgacctat 2400
ggtggacctc caaaaggaag gcgtctcccg tccaaccggt tctcgtcgaa tctcctagag 2460
ggttcctact ttcttacagt tcggtactac cttttacggg gaaggttagt cgacggaagt 2520
gttcgtccct agtctcgttg aggggccccc aggggttagg tgcgggaaga tcaggttggg 2580
gggagttact ctctccgtcc gtctagagtg ggtcgtgatc ctgtggtcct ccggtccctt 2640
tcgtagagac cgagtggtac attgtagacc gaacctcgtt cacccacaag acgtgtggtc 2700
cgtcgacgtg gagtgaccta gatcacaacg acgctcactg gagtgaagtc tcggggagat 2760
cgtctcaccc cgccttcagg actaccaacc acaggtactc caccttc 2807

<210> 30
<211> 2985
<212> DNA
<213> Human

<400> 30
cgcccggcgg tcacaccacc ttaagccgaa cagtgatcct cttgtaaaca caattaacgt 60
gacacgagac agttcctttg aaactaaata tcgaccccac gtgtttatta ccaacggcca 120
gcgtgtacct aagccatctt gaaacggaag gacttagaaa aagggacgtg atgctccttc 180
tcgaagagac gtcttacagt ttgtttctag ctgtgtaact aaggtcgaca agcaggaagt 240
agttctgcct tggaaggtcg ggtcggaggg actgcctgtc gcagttggtg gtgtcgggac 300
caccgagaag tctgcggtca ccctcgatgt caagttggta cttaccggta gtcttgcctg 360
aactgagcgg tggagagatg ggaagacgag gataggaccc tccatcaccc ggacagtcct 420
ttgacatact actgacgagg tcgtggtaac aacttctagg ggtctggttc acacttatgt 480
acgagttgag ctacgggttc tctgacacaa atcacacacc actgtagcga agacccatgg 540
tgatacccca tcgtagtaca cttcggacgt tccgtaagaa gttctcctgt taagttccgt 600
tatatcttat gtcgacggga cggtgcttac ttacacttta gtgtttcgcg tctgcattta 660
ggacggtccg aacggcgaag tacttcacaa attttcaccc gtacgacttt cttccccacg 720
cagaactgtc tcatgcacct ccagccgtct tcatgttcgc gtcctatcta cgcctcttgt 780
cgggtatgga cttgggagtc gaccaagtcg gtcggttttt cggtatattg ttctaacaga 840
gtgtaaacaa ccaccgactt ggcctcttct agatacggta cggactggga tgacaggggc 900
tgtcactgta gtttcgggag tgatgtgaca cactgaaccg gctggctctc aaccaccaat 960
agtaacctac ccgcttcgta taaggtccga agaggtgcga cagggaccgc ctggtctact 1020
cggaagacgt ctcacgaacc tacctttaaa actaggaacc acagcatatg gccagagaaa 1080
gtaaactcct acttgaacag atacgtctgc taatatatta cctgcttctg gtcaggttta 1140
atcgtccgga agaactagat ttattacgat aggacgtcga ccatttcttt atgttctcgt 1200
acttcgacct ttttcttctt aaacagtggg agtttcgata tcgagaacga ttaagtctga 1260
ggtacgtgta tcttctacaa cttcggcaag tcttcgaagt cctacagaat gtacttcgcg 1320
acgtcctaat acttcgaccg gtcgtgtacc ttctgggagc agctcgaccg ttctacgact 1380
actgtgacgg tgaggactcc gtctggagat ggttccggca cgtcgtaaag atgttgtagt 1440
ttgatcttcc gtttcagggt tacgtgtttg aaaaaaacct ttacaacctc cggttccaga 1500
ctgattttcg agggacccgg aagggtagga agtacaactt tttccctttt atttgggttc 1560
tcactacagc ttctttgaat ctcaaatcaa ttgttgtagt ttttagttgt ctgacgtgac 1620
tattaaatcg tcgttctgat acttcgtcga aagtctaagg aggtatccaa ggactactca 1680
agaaagatga aagaggtagt agaagaaagg agaaagaagg gtgtaaagag aaagagaaat 1740
aaaaaagagg aaaagaagaa agtggaggga ataaagaaac gaagaaagta aggatcaagg 1800
gtaagaggaa ataaaagaag ggcagacgga cggaagaaag aaaagaaatg gatgagagta 1860
aggagagaaa agagtaggaa ggggaaaaaa gatttaaact ttatcgaaat caaatttttt 1920
tttttaggag ggaaggggga aaggaaaggg aaagaaagga aaaagggaaa ggaaaaggga 1980
aaggaaagga aaggagaact ggaagaaagg tagaaagaaa aagaaggaag acgacgactt 2040
gaaaattttc tccagagatt gacttctctc taccttcggt cgggacggtt tcctacctct 2100
aggtattata cctacggtca cttgaataac acttggtatg gcaggggtta ctgattcctt 2160
agtttctctc tcttggttgc aaggattttc atgtcacgtt gtatatgttt aactgactca 2220

```
cgtcataatc taaagtaccc tcgtcggaga ttaatctgtt gaattcgttg caacgtagcc    2280
gacgaagaat agtaacgaaa aggtagatct agtcaatgtc ggtaaactaa ggaattaaca    2340
aaaaagttca gaaggtccat aaacaatcaa atcgatgata cattgaaaaa gtcccttatc    2400
aaattcgaaa taagtaagta cgttatgatt tctctttatt cttatgacgt taaaacacga    2460
ccgaaacttg ttaatgcttg ttattacttc ctgtttactt aggacttcct tctaaaaatt    2520
tttacaaaac aaagaagaat gtttacctct aaaaaaacat ggtcgaaatg gtgaaaagtc    2580
ggtaaataat tataccctta aattgaatga gttcgttatc aacttccctt ccacgtataa    2640
tagtgcctac gttaaataca acacacggtc agaccagggt ttgtagttaa agaattgtac    2700
tcgaggtcaa atggatttac aagtgactgt gtttcctact ctaatgtgga tgtcactgag    2760
actcatcagt gtatatattc gtgacgtgta ctctatatct aggcatctta acagtcctca    2820
cgtggagaga tgaaccctcc atgttaacgg tatactaaag atcgacggta ccaccaatcc    2880
ttacactatg acggacaaac gtttcagtgt ctggaacgga gtcttcctcg acactcggtc    2940
ataagtaaat tctcttaagg tggtgtgacc gccgggcgcg aacta                    2985
```